基于神经网络超参数优化方法的堆芯中子学参数预测研究

张凡 [1]　张俊达 [2]　孙启政 [2]　肖维 [2]　刘晓晶 [2]　张滕飞 [2]
1（上海交通大学 智慧能源创新学院　上海 200240）
2（上海交通大学 机械与动力工程学院　上海 200240）

摘要　神经网络可以基于大量数据学习输入输出变量之间的关系，具有强大的拟合能力，在包括核工程计算领域常用作程序的代理模型。中子输运计算作为中子学模拟的核心环节之一，其耗时较长的问题可以通过利用神经网络模型来解决。然而，神经网络模型具有一系列超参数需要设置，而手动调节这些超参数工作量大，重复繁琐，只能依靠经验进行，而且求解不同问题时这些超参数不可复用。为了解决以上问题，本文提出了一种采用贝叶斯优化(Bayesian Optimization)算法来调节神经网络超参数，结合了学习率衰减、损失函数优化方法，它可以针对不同问题的数据集，自动搜索超参数的最佳组合，以获得最佳性能，具有很高的灵活性和效率，泛化性强。本文对 TAKEDA 基准题得到的堆芯关键参数进行拟合，结果表明有效增殖因数 k_{eff} 的平均误差在 150pcm 以内，TAKEDA1 数据集上区域积分通量 ϕ 的平均误差率为 1.72%，最大误差率为 7.56%。该研究可为人工智能在堆芯物理计算理论的应用提供一定参考。

关键词　贝叶斯优化超参数；全连接神经网络；中子输运计算
中图分类号　TL329
DOI:

Research on Core Neutronic Parameter Prediction Based on Neural Network Hyperparameter Optimization Method

ZHANG Fan[1]　ZHANG Junda[2]　SUN Qizheng[2]　XIAO Wei[2]　LIU Xiaojing[2]　ZHANG Tengfei[2]
1(*College of Smart Energy, Shanghai Jiaotong University, Shanghai 200240, China*)
2(*School of Mechanical Engineering, Shanghai Jiaotong University, Shanghai 200240, China*)

Abstract　**[Background]**:Neural networks, with their powerful fitting capabilities, can learn the relationships between input and output variables based on large amounts of data, often serving as proxy models for physical programs in the field of engineering calculations, including nuclear engineering calculations. Neutron transport calculations, as one of the core links in neutronics simulations, often suffer from lengthy computational times. However, this issue can also be addressed by utilizing neural network models. Nevertheless, neural network models have a series of hyperparameters that need to be set, but manually adjusting these hyperparameters is laborious, repetitive, and reliant only on experience. Moreover, these hyperparameters are not reusable when solving different problems. [**Purpose**]: By seeking a surrogate model for VITAS, the research can provide some reference for the application of artificial intelligence in core physics calculation theory.**[Methods]**:This paper proposes the use of the bayesian optimization algorithm to adjust neural network hyperparameters, combined with learning rate decay and loss function optimization methods. **[Results]**: By fitting the key core parameters obtained from VITAS's calculation

---

基金名称项目：国家自然科学基金[12175138]，上海市青年科技启明星项目资助
第一作者：张凡，女，1999 年出生，2021 年毕业于山东大学，研究领域为机器学习方法在反应堆物理计算领域的应用
通讯作者：张滕飞，E-mail: zhangtengfei@sjtu.edu.cn
收稿日期：20XX-00-00，修回日期：20XX-00-00

of the TAKEDA benchmark problem, the results show that the average error of the effective multiplication factor is within 150pcm, and the average error rate of the regional integral flux on the TAKEDA1 dataset is 1.72%, with a maximum error rate of 7.56%. **[Conclusions]**: This approach can automatically search for the optimal combination of hyperparameters for different datasets to achieve the best performance, demonstrating high flexibility, efficiency, and strong generalization.

Key words Bayesian optimization for hyperparameter tuning, FCNN, Neutron transport computation

核反应堆物理计算中，蒙特卡罗方法[1]（Monte Carlo method，MC）和确定论方法是求解中子输运问题的两类主要方法。常用的中子输运方法包含有限元法（Finite Element Method，FEM）、特征线方法（Method Of Characteristics，MOC）和变分节块法[2-4]（Variational Nodal Method，VNM）等。这些传统中子输运方法存在计算效率低、耗时较长的特点，难以应用于频繁的堆芯方案优化、换料设计等工作。

相比于传统的蒙特卡罗方法和确定论中子输运方法，机器学习和人工智能在处理复杂大数据方面具有优势。近年来，核物理领域与机器学习的交叉应用正在兴起，具有广阔的应用前景。宋梅村等[5]开发出了基于 BP（back propagation）人工神经网络的堆芯参数预测程序,且相对误差在 10%以内,引起了后续研究者的兴趣，Akkoyun[6]使用人工神经网络实现对全聚变和聚变-蒸发反应截面的预测。而 Guo 等[7]则使用一个混合人工神经网络模拟红杉核电站的热力学行为。其中，全连接神经网络（Fully Connected Neural Network，FCNN），是一种连接方式较为简单的人工神经网络结构，属于前馈神经网络[8]的一种，只由输入层、隐藏层和输出层构成，并且在每个隐藏层中可以有多个神经元，可以应用于几乎所有任务的多功能学习方法，包括分类、回归，以及无监督学习。由于机器学习能够拟合复杂的分析计算过程并节省人工和计算的成本，因而获得了飞速的发展。

VITAS[9,10]是基于 VNM 开发的一套精确求解中子输运问题的通用型计算程序，它集成了多种计算方法，利用矩阵运算和数值积分技术，能够处理多种不同网格类型的多维、多群、稳态和瞬态中子输运问题。为改善传统输运方法计算效率低、耗时较长的缺陷，本文基于神经网络方法开发了堆芯输运计算替代模型，以替代 VITAS 程序的计算功能。

然而，神经网络中有一系列超参数需要设置，它们在极大程度上决定了 FCNN 的性能。手动调整超参数工作量庞大，需要一定的经验，代价高昂。而且随着机器学习的不断发展，深度学习（Deep Learning）建模中涉及到的超参数数量远超传统机器学习方法，手动调参开始不能满足精度和省时的需求。近年来贝叶斯优化[11]在求解黑盒函数问题中应用越来越广泛，已经成为超参数优化的主流方法[12]。贝叶斯优化在不同的领域也称作序贯克里金优化（Sequential Kriging Optimization，SKO）、基于模型的序贯优化（Sequential Model-based Optimization，SMBO）、高效全局优化（Efficient Global Optimization，EGO）。它是一种全局优化的方法，使用代理模型拟合真实目标函数，并根据拟合结果主动选择最有“潜力”的评估点进行评估，利用完整的历史信息来提高搜索效率，从而在少数次评估下得到复杂函数的最优解。因此，贝叶斯优化也称作主动优化（active optimization）。同时贝叶斯优化框架能够有效地利用完整的历史信息来提高搜索效率。

本文基于贝叶斯优化方法，结合学习率衰减、损失函数优化等方法，对 FCNN 超参数进行优化选取，使用 VITAS 程序计算 TAKEDA 基准题得到的实验数据作为数据集，评估了优化超参数后的 FCNN 的预测精度，并与手动调参的网络计算结果做了对比分析，旨在提供更高效求解堆芯换料优化问题的方法，为人工智能在核工业领域的进一步应用做出探索。

1 理论模型

1.1 贝叶斯优化方法

贝叶斯优化框架有两个关键部分：（1）使用概率模型代理原始评估代价高昂的复杂目标函数；（2）用代理模型的后验信息构造主动选择策略，即采集函数。在实际应用中，需要针对具体问题选择合适的模型。本节介绍超参数优化问题的定义和贝叶斯优化的步骤。在本节最后，给出了本研究选取的 FCNN 超参

数及其取值范围。

超参数优化问题定义为：

$$\mathrm{argmin}_{x\in X}\ f(x) \tag{1}$$

其中，x 为超参数的一组设置取值，X 为超参数搜索空间。$f(x)$ 为超参数优化中，需要优化的目标。本研究取损失函数（Loss Function）作为 $f(x)$ 的取值，对损失函数的比较和选取在 1.2 节（2）中损失函数优化里给出。在超参数算法优化的目标是以最快的方式找到全局最优解。贝叶斯优化超参数的步骤如下：

（1）在超参数搜索空间 X 中随机初始化 n_0 组 x_{init}；

（2）获得其对应的函数值 $f(x_{init})$，得到初始点集 $D_0=\{x_{init}, f(x_{init})\}$，令 $t=n_0, D_{t-1}=D_0$；

（3）根据当前获得的点集分布，构建代理模型 $g(x)$；

（4）基于代理模型 $g(x)$,最大化采集函数 $\alpha(x\,|\,D_{t-1})$,获得下一个评估点：$x_t=\arg\max\alpha(x\,|\,D_{t-1})$；

（5）获得评估点 x_t 的函数值 $f(x_t)$，将其加入到当前评估点合集中：$D_t=D_{t-1}\cup\{x_t, f(x_t)\}$，$t=t+1$。若 t 小于最大迭代次数 N，转步骤（3）；

（6）达到最大迭代次数后，输出最优评估点 $\{x^*, f(x^*)\}$。

本研究调用 Tensorflow[13]的 Bayesian Optimization 函数实现贝叶斯优化过程。通过输入超参数及其取值范围启动优化功能，选取的 FCNN 超参数及其取值范围如表 1 所示：

表 1　选取的超参数及其取值范围

Table 1　The selected hyperparameters and their ranges

超参数 Hyperparameters	取值范围 Ranges
批次大小 batch_size	[1,1000]
隐藏层数量 num_hidden_layers	[1, 5]
阈值 min_delta	[1e-6,1e-4]
衰减因子 factor	[0.1,0.9]
隐藏层神经元数量 num_neurons	[1, 2000]
损失函数参数 loss_delta	[0.1,10]

表中列出的超参数皆为 FCNN 中的关键参数，在下文中介绍。

1.2　全连接神经网络

在中子输运求解的代理模型选择方面，本工作选取全连接层神经网络 FCNN 开展研究。FCNN 是一种连接方式较为简单的人工神经网络结构，属于前馈神经网络的一种，只由输入层（Input Layer）、隐藏层（Hidden Layer）和输出层（Output Layer）构成，并且在每个隐藏层中可以有多个神经元，如图 1 所示。

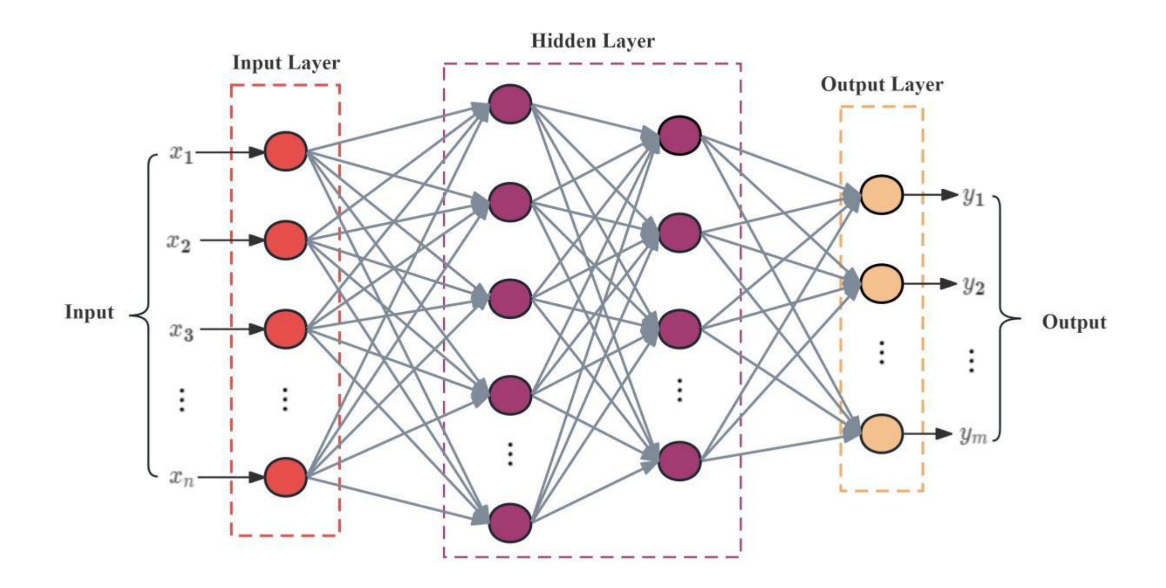


图 1　FCNN 示意图

Fig.1　Schematic of FCNN

整个网络的搭建基于 python 上的开源框架 Tensorflow，它是 2015 年由谷歌公司研发的第二代人工智能深度学习平台。FCNN 训练过程中有如下参数：

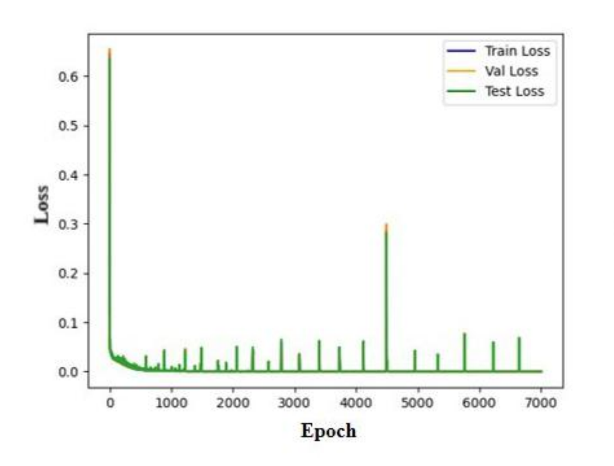
Epoch: 将全部的训练集数据投入神经网络模型完成一次完整的训练，该过程称为一个 Epoch。

Batch_Size（一次投入网络模型中的数据量）：因为算力或者其他原因，不能将数据一次性通过神经网络时，就需要将数据集分成若干个 batch（批次），使用训练集中的一小部分数据样本对模型权重进行一次反向传播的参数更新。

学习率（Learning Rate，lr）是指在训练神经网络时用于调整拟合参数的步进大小，它决定了每次梯度更新时参数的调整步长。学习率的选择直接关系到模型的性能和训练过程的效果。

（1）自适应调整学习率

学习率若设置的过大，会导致步长变大，容易出现震荡，导致不易找到高精度的解。图 2（左）所示为过大学习率导致的损失函数震荡，结果不易稳定的收敛。学习率若设置的太小，收敛速度会变慢，容易只得到局部最优解。基于此，本文提出一种自适应调整学习率的方法，刚开始的时候用较大的学习率进行快速逼近，然后逐渐降低学习率，以适应模型参数更新的规律，消除后期损失函数 loss 的震荡，loss 曲线趋于稳定。图 2（右）所示为自适应调整学习率后的损失函数下降曲线，很好地消除了震荡。



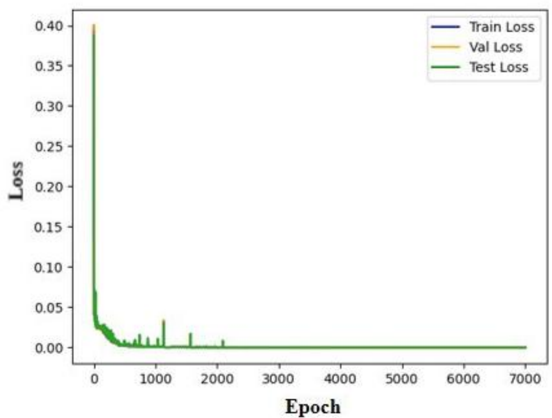


图 2 不同学习率下损失函数曲线图

Fig.2 Plots of loss function curves with different learning rates

本研究使用的 ReduceLROnPlateau 函数是 tensorflow 中的一个学习率调度器，它能够根据监测指标的变化自动调整学习率。当验证集上的性能指标停止改善时，ReduceLROnPlateau 会逐渐减小学习率，以便模型更好地收敛，它有如下几个参数：

lr： 学习率初始值；

factor：每次减少学习率的因子，学习率将以 $\mathrm{lr} = \mathrm{lr} * \mathrm{factor}$ 的形式被减少；

patience：当 patience 个 epoch 过去而模型性能不提升时，学习率减少的动作会被触发；

min_delta：阈值，模型性能指标改善值小于 min_delta 时，视为性能没有提升。

其中，factor, min_delta 这两个参数被选取为本研究中贝叶斯优化的超参数。

（2）损失函数优化

均方误差（Mean Square Error，MSE），指的是模型预测值 $f(x)$ 与样本真实值 y 之间距离平方的平均值。其公式如式（2）所示：

$$\mathrm{MSE} = \frac{1}{m}\sum_{i=1}^{m}(y_i - f(x_i))^2 \tag{2}$$

平均绝对误差（Mean Absolute Error，MAE）指的就是模型预测值 $f(x)$ 与样本真实值 y 之间距离的平均值。其公式如式（3）所示：

$$\mathrm{MAE} = \frac{1}{m}\sum_{i=1}^{m}|y_i - f(x_i)| \tag{3}$$

公式（2）、（3）中，y_i 和 $f(x_i)$ 分别表示第 i 个样本的真实值和预测值，m 为样本个数。

使用 MAE 作为损失函数训练神经网络的一个大问题是其梯度始终很大，可能导致使用梯度下降训练模型结束时遗漏最小值。而 MSE 的梯度会随着损失值接近其最小值逐渐减少，从而使其更准确。该问题可以选择使用 Huber Loss 来进行改善。Huber Loss 是回归问题中的一种损失函数，它结合了 MSE 和 MAE 两者的特点。它可以在最小值附近减小梯度，解决 MAE 漏掉最小值的问题；而且相比于 MSE，它对异常值更具鲁棒性。本文中使用 Huber Loss 作为损失函数，其公式如式（4）所示：

$$L_\delta(y, f(x)) = \begin{cases} \frac{1}{2}(y - f(x))^2, |y - f(x)| \le \delta \\ \delta|y - f(x)| - \frac{1}{2}\delta^2, |y - f(x)| > \delta \end{cases} \tag{4}$$

Huber Loss 是对 MSE 和 MAE 的综合，它包含了一个超参数 δ ，δ 值的大小决定了 Huber Loss 对 MSE 和 MAE 的侧重性，当 $\delta \to 0$ 时，Huber Loss 会趋向于 MAE；当 $\delta \to \infty$ 时，Huber Loss 会趋向于 MSE。在表 1 中，δ 就是 loss_delta。

上述参数的设置决定了 FCNN 的性能，于是把它们设置为贝叶斯优化的超参数，在表 1 中给出了它们的搜索空间。

1.3 数据集的获取

本文使用的数据集通过 VITAS 程序计算 TAKEDA1、2 基准题得出，分别为 10000 与 20000 组，每组数据样本由堆芯排布方式与两个堆芯参数（有效增殖因数 k_{eff} 和 ϕ 区域积分通量）组成，输入维度由反应堆堆芯排布决定，输出维度为堆芯参数。TAKEDA1 为 1/8 对称轻水反应堆模型，TAKEDA2 为 1/4 快中子增殖反应堆模型。

TAKEDA1 包含三种组件，分别为控制棒（Control Rod，CR）、反射层（Reflector）和燃料（Core），它的三维示意图和 *x-y* 截面堆芯排布如图 3 所示。为了保证堆芯有物理意义，将堆芯 *x-y* 截面的 1/4 最外一层的反射层固定位置，其余组件打乱顺序，再对称成完整的堆芯，如图 4。为了便于 FCNN 计算，将三种组件 CR，Reflector，Core 分别映射为-1， 0，1。如此生成 10000 组不同的随机排布，每个输入向量中-1，0，1 的排列顺序与其组件打乱后的排布顺序一一对应，输入维度为 16，训练时分别用 2 套独立的 FCNN 预测 2 个物理量，k_{eff} 的输出维度为 1，ϕ 为 6。

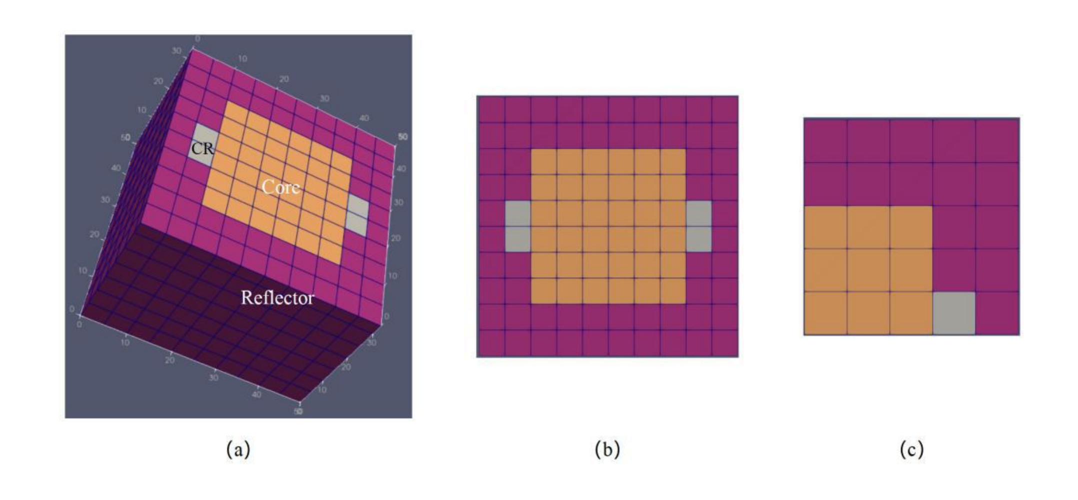


图 3 TAKEDA1 示意图：(a) 三维模型 (b) *x-y* 截面 (c) *x-y* 截面的 1/4

Fig.3 Schematic of TAKEDA1:

(a) 3D model, (b) *x-y* cross-section, (c) Quarter of the *x-y* cross-section.

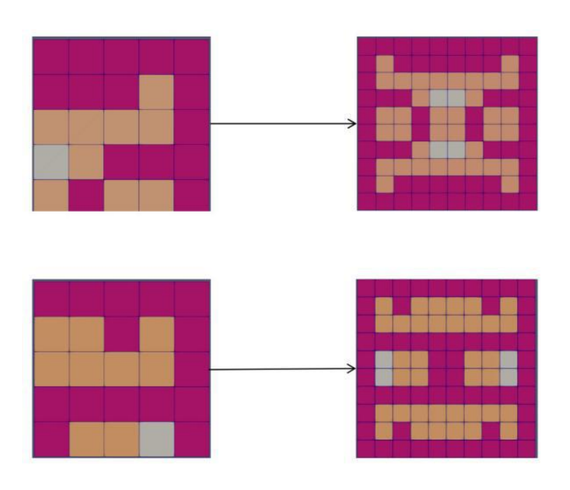


图 4 TAKEDA1 堆芯排布打乱示意图

Fig.4 TAKEDA1 core arrangement disruption schematic

TAKEDA2 包含三种组件，分别为控制棒（Control Rod，CR），轴向增殖区（Axial Blanket）和径向增殖区（Radial Blanket），它的示意图如图 5。为了保证堆芯有物理意义，将堆芯 *x-y* 截面的 1/4 最外三层 Radial Blanket 固定位置，其余组件打乱顺序，再对称成完整的堆芯，如图 6。为了便于 FCNN 计算，将三种组件 CR，Axial Blanket，Radial Blanket 分别映射为-1， 0，1。如此生成 20000 组不同的随机排布。每个输入向量中-1，0，1 的排列顺序与其组件打乱后的排布顺序一一对应，输入维度为 121，训练时分别用 2 套独立的 FCNN 预测 2 个物理量，k_{eff} 的输出维度为 1，ϕ 为 20。

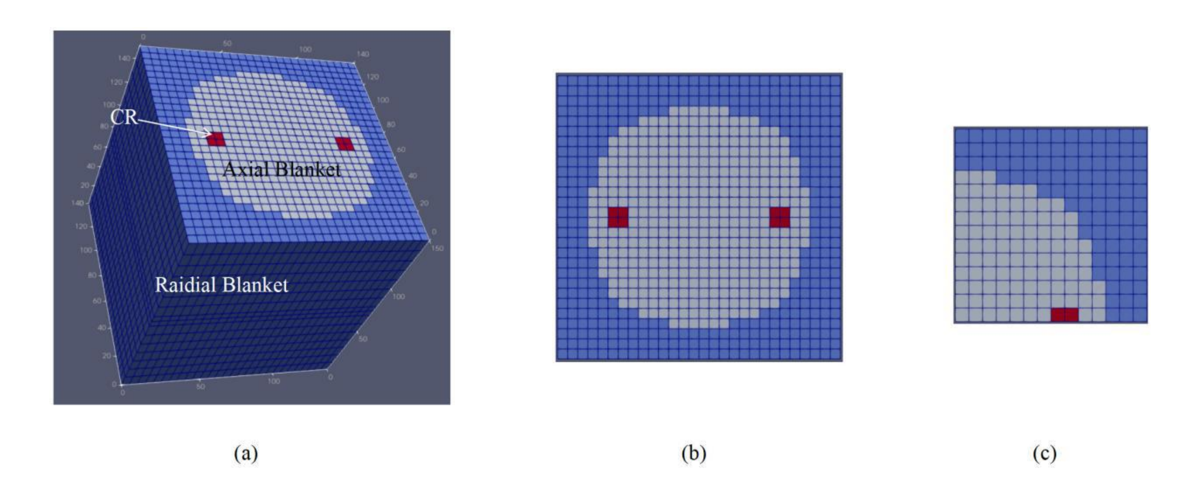


图 5 TAKEDA2 示意图：(a) 三维模型 (b) *x-y* 截面 (c) *x-y* 截面的 1/4

Fig.5 Schematic of TAKEDA2:

(a) 3D model, (b) *x-y* cross-section, (c) Quarter of the *x-y* cross-section.

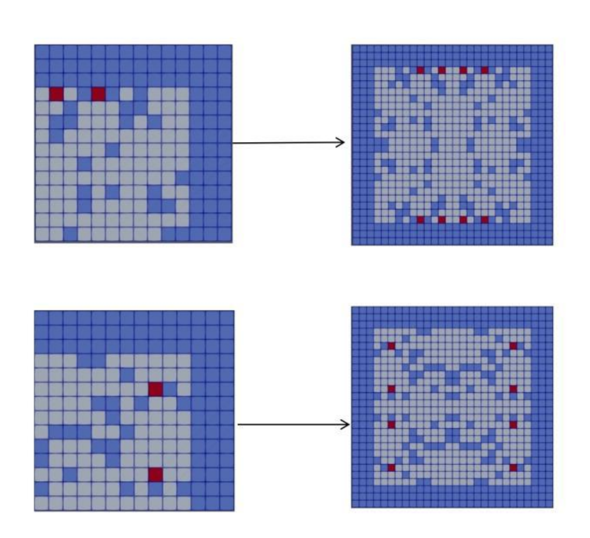


图 6 TAKEDA2 堆芯排布打乱示意图

Fig.6 TAKEDA2 core arrangement disruption schematic diagram

训练 FCNN 时需要将原始数据转化为标准数据集，假设有 N 个训练样本，每个训练样本应为：

$$T_k = (X_k, Y_k)(k = 0,1,\cdots,N-1) \tag{5}$$

T_k 为第 k 个训练样本，包含第 k 个输入变量 X_k 和输出变量 Y_k，其中，

$$X_k = [x_{k1}, x_{k2}, \cdots, x_{ki}] \tag{6}$$

$$Y_k = [y_{k1}, y_{k2}, \cdots, y_{ki}] \tag{7}$$

i，j 分别为输入、输出变量的维度。TAKEDA1 中的 k =10000,输入向量 X 为-1，0，1 的排列顺序，每组输入的维度 i =16，分别用 2 套独立的 FCNN 预测 2 个物理量，两套 FCNN 的输出变量 Y 分别为 k_{eff} 和 ϕ，k_{eff} 的维度 j =1，ϕ 为 6。TAKEDA2 中的 k =20000,输入向量 X 为-1，0，1 的排列顺序，每组输入的维度 i =121，分别用 2 套 FCNN 的网络预测 2 个物理量，两套 FCNN 的输出变量 Y 分别为 k_{eff} 和 ϕ，k_{eff} 的维度 j =1，ϕ 为 20。

1.4 模型验证

本研究的数据集使用了 10000 组 TAKEDA1 和 20000 组 TAKEDA2 数据，输入为反应堆的堆芯排布，输出的两个参数分别为 k_{eff} 和 ϕ，以 6:4 的比例划分为训练集和验证集，算法流程图如图 7 所示：

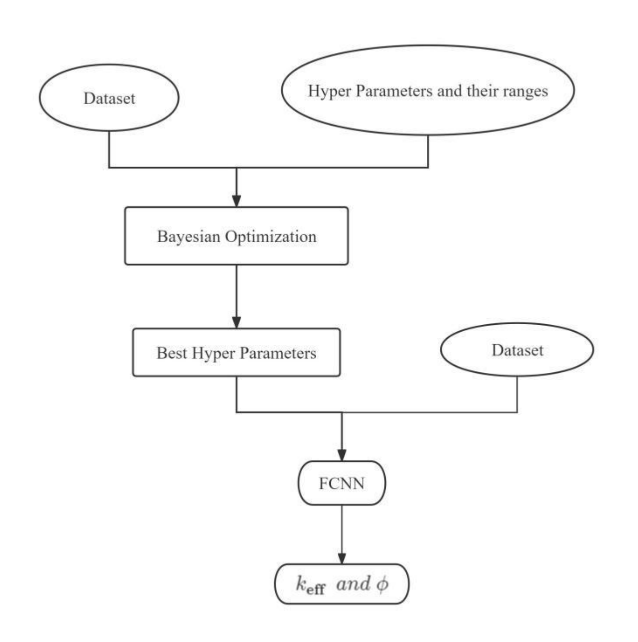


图 7 算法流程图

Fig.7 Algorithm flowchart

以 TAKEDA1 为例，下表给出了手动设置的一组超参数以及贝叶斯优化输出的一组超参数的值：

表 2　TAKEDA1 的超参数选择

Table 2　Hyperparameter selection for TAKEDA1

		批次大小 Batch_size	隐藏层数 Hidden_layers	阈值 Min_delta	衰减因子 Factor	神经元数量 Num_neurons	损失函数参数 Loss_delta
手动设置 Manual Setting	k_{eff}	2000	4	5.5e-5	0.9	800, 900, 1600, 800	0.608
	ϕ	3	5	4.5e-5	0.5	400, 800, 1800, 600, 50	0.3
贝叶斯优化 Bayesian Optimization	k_{eff}	53	3	1.248e-5	0.5	785, 1683, 1042	9.6
	ϕ	11	5	8.927e-05	0.61	1033, 1741, 237, 1937, 579	7.65

代入 FCNN 训练后，TAKEDA1 两个预测参数模型的训练过程曲线在下面给出：

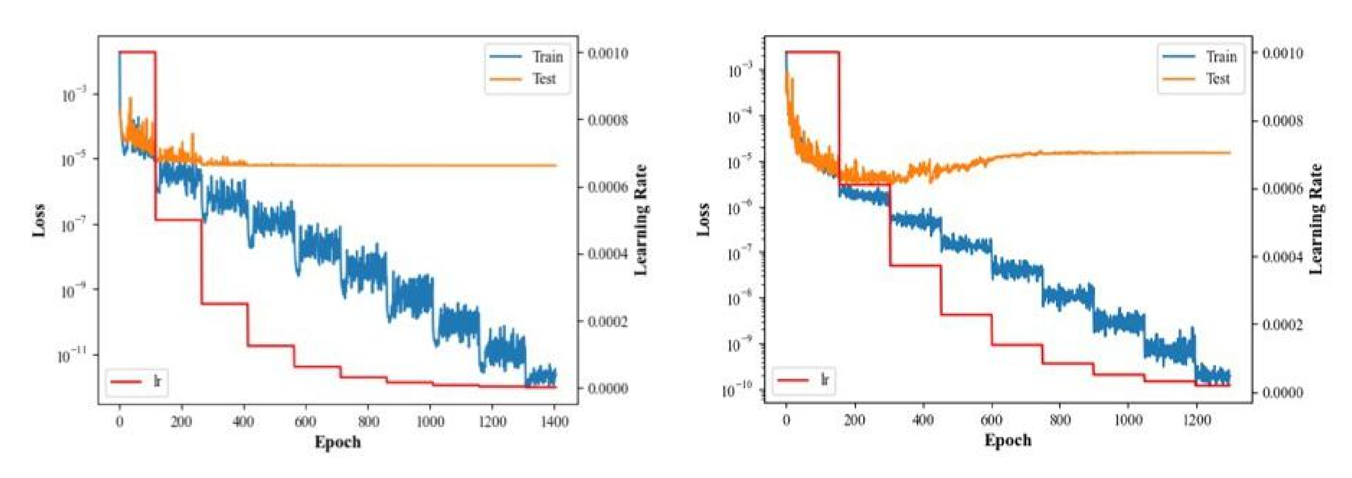


图 8　k_{eff} 的贝叶斯优化超参数（左）与手动调参模型（右）损失函数曲线的比较

Fig.8　Comparison of loss function curves between bayesian optimization and manual tuning of k_{eff}

从图 8 中可以看出，随着学习率的下降，训练集的损失函数在逐步下降，而图 8（左）的测试集损失函数在下降了 400 个 epoch 后趋于平稳，图 8（右）手动调参组的测试集损失函数先下降后缓慢上升，最后趋于平稳。

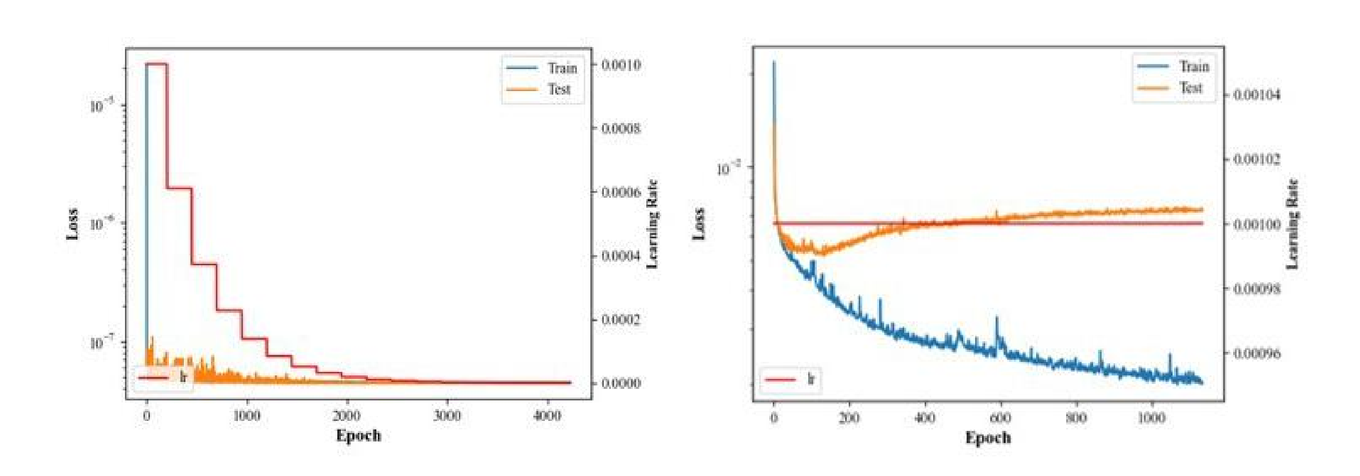


图 9 ϕ 的贝叶斯优化超参数（左）与手动调参模型（右）损失函数曲线的比较

Fig.9 Comparison of loss function curves between bayesian optimization and manual tuning of ϕ

图 9（左）贝叶斯调参组的损失函数在刚开始训练时就迅速下降，而图 9（右）手动调参组的训练集损失函数下降时，测试集的损失函数反而在上升，出现了过拟合。

2 结果和讨论

预测问题评价指标：MAE , MSE 的定义在式(2)(3)中给出，平均百分比误差（Mean Absolute Percentage Error，MAPE）和 R^2（拟合优度）的定义为：

$$\text{MAPE}=\frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i-y_i}{y_i}\right| \tag{8}$$

$$R^2=1-\frac{\sum(y_i-y_i)^2}{\sum(y_i-\overline{y_i})^2} \tag{9}$$

y_i 表示真实值，$\overline{y_i}$ 表示样本均值，y_i 表示预测值。

R^2 用于度量因变量的变异中可由自变量解释部分所占的比例，取值范围是 0~1，R^2 越接近 1，表明回归曲线与各观测点越接近，回归的拟合程度就越好。

表 3 和表 4 分别是有效增殖因数 k_{eff} 和区域积分通量 ϕ 这两个输出变量在贝叶斯优调参后 FCNN 输出结果和手动设置后输出结果误差的比较，以 TAKEDA1 和 TAKEDA2 作为数据集：

表 3 贝叶斯调参与手动调参模型预测 k_{eff} 的误差比较

Table 3 Comparison of errors between bayesian optimization and manual hyperparameter tuning for k_{eff}

		平均绝对误差 MAE	均方误差 MSE	拟合优度 R^2
TAKEDA1	手动设置 Manual Setting	397pcm	6.78996e-05	0.9981
	贝叶斯优化 Bayesian Optimization	118pcm	0.000003	0.9997
TAKEDA2	手动设置 Manual Setting	1625pcm	0.07219	0.2756
	贝叶斯优化 Bayesian Optimization	132pcm	0.000005	0.9988

从表 3 中可以看出，贝叶斯优化的超参数组合与手动设置的超参数在误差大小和拟合优度方面相比优势明显。贝叶斯优化得到的 TAKEDA1 的 k_{eff} 平均绝对误差为 118pcm，TAKEDA2 的 k_{eff} 平均绝对误差为 132pcm，与手动调参得到的 397pcm 和 1625pcm 相比，在 TAKEDA1 数据集上，贝叶斯优化的平均绝对误差为手动调参的 1/3；在 TAKEDA2 数据集上，贝叶斯优化的误差为手动调参的 1/12，精度得到了明显提升。贝叶斯优化得到的 R^2 在 0.999 左右，手动调参得到的 R^2 在贝叶斯优化之下，表明贝叶斯优化后 FCNN 的拟合程度更好。

表 4 贝叶斯调参与手动调参模型预测 ϕ 的误差比较

Table 4　Comparison of errors between bayesian optimization and manual hyperparameter tuning for ϕ

		平均百分比误差/% MAPE/%	最大百分比误差/% Max Percentage Error/%	均方误差 MSE	拟合优度 R^2
TAKEDA1	手动设置 Manual Setting	14.8305	365.3688	0.0000000583	0.8469
	贝叶斯优化 Bayesian Optimization	1.7188	7.5586	0.000000000234	0.9818
TAKEDA2	手动设置 Manual Setting	4.4897	92.1902	4.258e-11	0.9907
	贝叶斯优化 Bayesian Optimization	0.8214	11.1301	3.64e-12	0.9939

表 4 中贝叶斯优化得到的 TAKEDA1 的ϕ平均百分比误差为为 1.7187%，最大百分比误差为 7.5585%，TAKEDA2 的ϕ平均误差为 0.8213%,最大百分比误差为 11.130%，达到了一定的精度。与手动调参相比，在 TAKEDA1 数据集上，贝叶斯优化的平均百分比误差为手动调参的 1/9，最大百分比误差为手动调参的 1/50；在 TAKEDA2 数据集上，贝叶斯优化的平均百分比误差为手动调参的 1/5，最大百分比误差为手动调参的 1/8，精度同样得到了很大提升。贝叶斯优化得到的 R^2 在 0.99 左右，手动调参得到的 R^2 在贝叶斯优化之下，表明贝叶斯优化后 FCNN 的拟合程度更好。图 10, 11 分别为$k_{\rm eff}$和ϕ经过贝叶斯方法调参后输出结果的误差分布直方图：

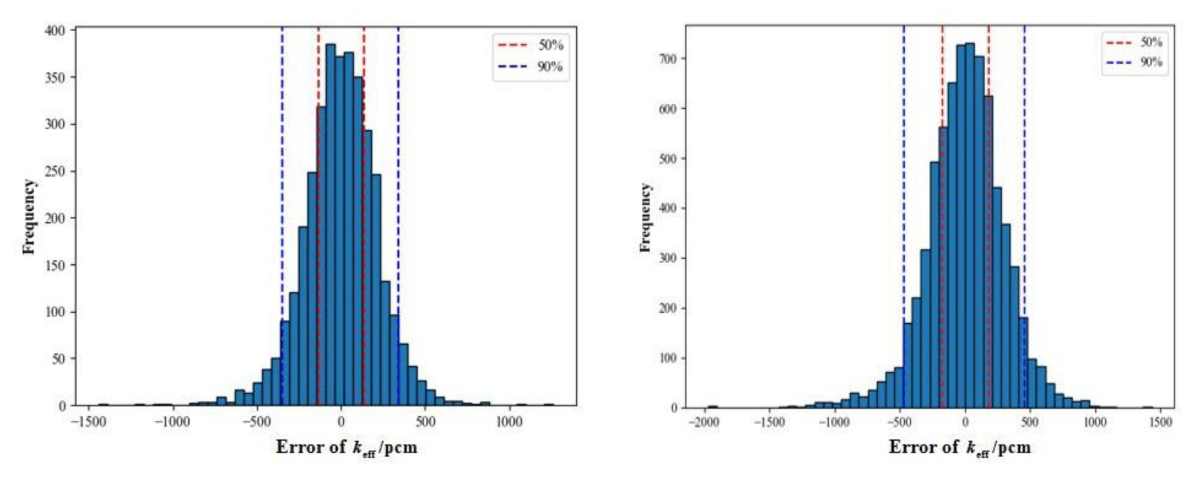
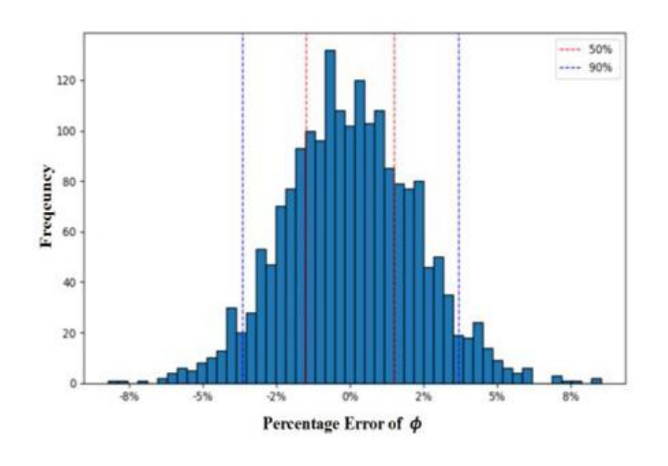


图 10　$k_{\rm eff}$ 的误差分布直方图

Fig.10　Error distribution of $k_{\rm eff}$

由图 10 可以看出，TAKEDA1 和 TAKEDA2 中$k_{\rm eff}$的误差呈正态分布，而且分布都较集中，90%的误差都集中在 500pcm 以内，50%的误差都集中在 150pcm 以内。



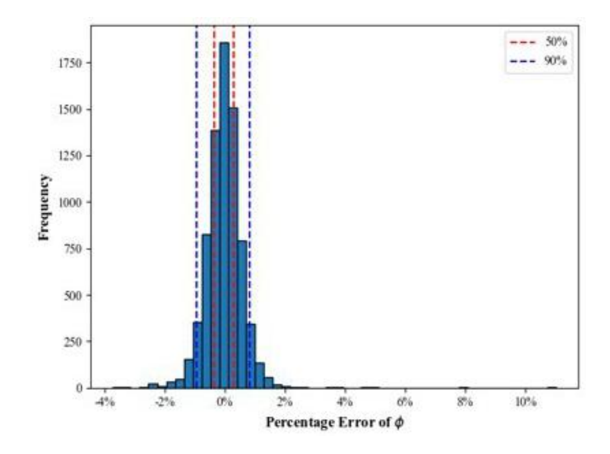


图 11 ϕ的百分比误差分布直方图

Fig.11 Percentage error distribution of ϕ

由图 11 可以看出，TAKEDA1 和 TAKEDA2 中ϕ的百分比误差都呈正态分布，TAKEDA1 的误差较分散，TAKEDA2 的图像“瘦高”，误差分布比较集中，TAKEDA1 中ϕ 90%的百分比误差都集中在 3.5%以内，TAKEDA2 中ϕ 90%的百分比误差都集中在 1%以内，达到了良好的精度。

3 结语

为更高效求解堆芯换料优化问题，本文提出了贝叶斯优化神经网络超参数算法作为堆芯输运计算替代模型，用 TAKEDA1, TAKEDA2 基准题获取的数据集针对进行训练并与手动调参的结果进行对比分析，研究表明：

（1）贝叶斯优化超参数后的 FCNN 能够较好地拟合 VITAS 代码的计算结果，贝叶斯优化得到的k_{eff}的平均误差在 150pcm 以内，TAKEDA1 的ϕ平均百分比误差为为 1.7187%，TAKEDA2 的ϕ平均百分比误差为 0.8213%，误差在允许范围内。

（2）贝叶斯优化得到的超参数建立的 FCNN 模型明显优于手动调参的模型，在训练得到的误差和拟合优度两个层面上来讲都是如此。

（3）本文使用的超参数及对应的 FCNN 只适用于 VITAS 计算得到的 TAKEDA1 和 TAKEDA2 数据集。若要对其他堆芯参数进行预测，可以替换对应的数据集，按照图 7 所示的流程图得到新的超参数组合并训练 FCNN 即可，可以代替手动调参的步骤，具有普适性。

（4）验证了神经网络预测参数方法在反应堆物理计算领域的可行性和优势，对于不同种类的神经网络在反应堆物理计算领域的应用，还需进一步探索更多的可能性。

作者贡献声明 张凡负责模拟计算以及论文撰写；张俊达、孙启政、肖维负责提供技术支持；刘晓晶负责基金支持；张滕飞负责基金支持、文章修改。

参考文献

1 Kazuteru SUGINO,Toshikazu TAKEDA.An Improvement of the Transverse Leakage Treatment for the Nodal SN Transport Calculation Method in Hexagonal-Z Geometry[J].Journal of Nuclear Science and Technology,1996(8).DOI:10.3327/jnst.33.620.

2 PALMIOTTI G, LEWIS E E, CARRICO C B. VARIANT: VARIational Anisotropic Nodal Transport for multidimensional Cartesian and hexagonal geometry calculation: ANL-95/40[R]. Argonne, IL, USA: Argonne National Laboratory, 1995.

3 ZHANG T F, WU H C, CAO L Z, et al. An improved variational nodal method for the solution of the three dimensional steady-state multi-group neutron transport equation[J]. Nuclear Engineering and Design, 2018, 337: 419-427.DOI:10.1016/j.nucengdes.2018.07.009.

4 ZHANG T F, ZHU L, WU H C, et al. Development and validation of high flux engineering test reactor fuel management platform HEFT[C]. Atomic Energy Science and Technology, 2013, 47(S1): 467-471.

5 宋梅村，蔡琦．基于 BP 神经网络的反应堆功率预测[J]. 原子能科学技术,2011,45(10):1242-1246.DOI:10.7538/yzk.2011.45.10.1242.

SONG Meicun,CAI Qi.Reactor power prediction based on BPNeuralNetwork[J].Atonomic Energy Science and Technology,2011,45(10):1242-1246.DOI:10.7538/yzk.2011.45.10.1242.

6 AKKOYUN S. Estimation of fusion reaction cross-sections by artificial neural networks[J]. Nuclear Instruments and Methods in Physics Research B, 2020, 462: 51-54. DOI:10.1016/j.nimb.2019.11.014.

7 GUO Z, UHRIG R E. Use of artificial neural networks to analyze nuclear power plant performance[J].Nuclear Technology, 1992, 99(1): 36-42.DOI:10.13182/NT92-A34701.

8 RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors[J].Nature, 1986, 323: 533-536.DOI:10.1038/323533a0.

9 ZHANG T F, XIAO W, YIN H, et al. VITAS: A multi-purpose simulation code for the solution of neutron transport problems based on variational nodal methods.[J].Annals of Nuclear Energy ,2022.DOI:10.1016/j.anucene.2022.109335.

10 张滕飞，殷晗，孙启政，等．通用型中子输运程序 VITAS 应用研究[J]. 核动力工程,2023,44(02):15-23.DOI:10.13832/j.jnpe.2023.02.0015.

Zhang Tengfei, Yin Han, Sun Qizheng, Xiao Wei, et al.Application Research on VITAS—a General-purpose Neutron Transport Code [J].Nuclear Power Engineering,2023,44(02):15-23.DOI:10.13832/j.jnpe.2023.02.0015.

11 崔佳旭,杨博.贝叶斯优化方法和应用综述[J].软件学报,2018,29(10):3068-3090.DOI:10.13328/j.cnki.jos.005607.

CUI Jiaxu, YANG Bo.Survey on Bayesian Optimization Methodology and Applications[J].Journal of Software,2018,29(10):3068-3090.DOI:10.13328/j.cnki.jos.005607.

12 Ghahramani Z. Probabilistic machine learning and artificial intelligence [J].Nature, 2015,521:452-459.DOI:10.1038/nature14541.

13 MARTIN A, BARHAM P, CHEN J, et al.TensorFlow: A system for large-scale machine learning[J].USENIX Association, 2016.DOI:10.48550/arXiv.1605.08695.